



The University of Adelaide
School of Applied Mathematics
Faculty of Engineering, Computer and Mathematical Sciences

Vido Gluščević

Approximation for Csiszár f -divergence

Thesis submitted for the degree of
Doctor of Philosophy
at The University of Adelaide

Adelaide, March, 2004.

CONTENTS

Signed statement	i
Acknowledgement	ii
Abstract	iii
1 Introduction	2
2 f-divergence	4
2.1 Introduction	4
2.2 Properties of f -divergence	4
2.3 Particular instances of Csiszár f -divergence measures	6
2.4 Some inequalities for f -divergence measures	8
3 Shannon entropy and mutual information	9
3.1 Shannon entropy	9
3.1.1 Introduction	9
3.1.2 Properties of Shannon entropy	9
3.1.3 Inequalities related to Shannon's entropy	13
3.2 Mutual information	16
4 Inequalities required	17
4.1 Ostrowski integral inequality for an absolutely continuous function	17
4.2 Ostrowski type inequality for mappings of bounded variation	19
4.3 Diaz and Metcalf inequality	21
4.4 Iyengar inequality	21
4.5 Trapezoid inequality	22
4.6 Generalised triangle inequality	22
5 Approximation for Csiszár f-divergence	23
5.1 Introduction	23
5.2 Csiszár f -divergence, Ostrowski's inequality and mutual information	23
5.2.1 Introduction	23
5.2.2 An inequality for Csiszár f -divergence	24
5.2.3 Applications to some common divergence measures	27
5.2.4 Some numerical examples	37

5.2.5	Application to mutual information	37
5.3	Approximations for Csiszár f -divergence via midpoint inequalities . .	38
5.3.1	Introduction	38
5.3.2	Preliminaries	39
5.3.3	Basic results	43
5.3.4	Applications to some common divergence measures	46
5.3.5	Application to mutual information	54
5.4	The approximations for Csiszár f -divergence for absolutely continuous mappings	54
5.4.1	Introduction	54
5.4.2	Basic theorems	54
5.4.3	Applications to some common divergence measures	57
5.5	The approximation of Csiszár f -divergence for mappings of bounded variation	64
5.5.1	Introduction	64
5.5.2	Main results	65
5.5.3	Application to some common divergence measures	67
6	References	74

LIST OF FIGURES

1	Kullback-Leibler distance - Theorem 1	32
2	Negative Kullback-Leibler distance - Theorem 1	33
3	Hellinger discrimination	34
4	Variational distance - Theorem 1	35
5	Triangular discrimination - Theorem 1	36
6	Kullback-Leibler distance - Theorem 2 and 3	50
7	Negative Kullback-Leibler distance - Theorem 2 and 3	51
8	Hellinger discrimination - Theorem 2 and 3	52
9	Renyi distance - Theorem 2 and 3	53
10	Kullback-Leibler distance - Theorem 4 and 5	61
11	Negative Kullback-Leibler distance - Theorem 4 and 5	62
12	Helinger discrimination - Theorem 4 and 5	63
13	Kullback-Leibler distance - Theorem 6	72
14	Variational distance - Theorem 6	73

SIGNED

This work contains no material which has been accepted for the award of any other degree or diploma in any University or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

SIGNED.....*Vicki Glasstone*.....*12/8/2004*.....DATE

ACKNOWLEDGEMENT

My deepest appreciation go to my supervisors, Prof. Charles Pearce and Prof. Sever Dragomir, whose guidance, encouragement and intellectual support have sustained me throughout creation of this thesis.

This doctorate would not have been possible without the continued support of my wife Jadranka and my daughter Matea.

SUMMARY

A variety of divergences have been introduced in the literature for measuring the difference between two probability distributions. A unifying concept is the Csiszár f -divergence, which subsumes many of these as special cases.

In many applications, we have only limited information about the distributions concerned. In this thesis we address this issue in the case of distributions with finite support. We make substantial use of the theory of inequalities, in particular, the Diaz-Metcalf inequality, Iyengar's inequality and the trapezoidal inequality.

Our main results are approximations for Csiszár f -divergence, with application to mutual information. Special cases for various choices of the function f are also of interest. A number of our results are considerable improvements on existing results, in one case being eight times sharper than the best pre-existing result.

Some of the material, in the thesis has been published in seven papers [38]-[44].